

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Stres, sebagai sebuah respons psikofisiologis yang kompleks, muncul ketika individu dihadapkan pada tuntutan lingkungan, baik emosional maupun fisik, yang melampaui kapasitas adaptif mereka [1], [2]. Fenomena ini bukan sekadar ketidaknyamanan sesaat, melainkan sebuah kondisi yang dapat memicu serangkaian perubahan biologis dan psikologis dalam tubuh. Paparan stres yang berkepanjangan, atau stres kronis, telah diidentifikasi secara luas dalam literatur ilmiah sebagai faktor risiko signifikan yang berkontribusi terhadap patogenesis berbagai kondisi kesehatan merugikan. Beberapa di antaranya meliputi gangguan sistem kardiovaskular seperti hipertensi dan penyakit jantung koroner, serta gangguan pada ranah kesehatan mental seperti gangguan kecemasan generalisata, depresi, dan gangguan tidur[3]. Mengingat dampak multifaset dan potensi debilitas yang ditimbulkannya terhadap kualitas hidup dan produktivitas individu, pengembangan metode deteksi dini tingkat stres menjadi suatu urgensi. Intervensi yang dilakukan pada tahap awal memiliki potensi besar untuk mencegah eskalasi stres menjadi kondisi kronis dan memitigasi risiko komplikasi penyakit yang lebih serius, sehingga memungkinkan penanganan yang lebih proaktif dan efektif[1], [4].

Seiring dengan kemajuan teknologi sensor dan pemrosesan sinyal, pendekatan objektif untuk asesmen stres melalui analisis sinyal fisiologis tubuh semakin mendapat perhatian. Di antara berbagai modalitas sinyal fisiologis, sinyal elektrokardiogram (EKG) menonjol sebagai salah satu kandidat yang paling menjanjikan[5], [6]. EKG, yang secara fundamental merekam aktivitas kelistrikan jantung, menyediakan jendela non-invasif ke dalam fungsi sistem saraf otonom. Perubahan dalam pola EKG, terutama pada parameter seperti variabilitas denyut jantung (Heart Rate Variability - HRV), diketahui sangat sensitif terhadap fluktuasi kondisi psikofisiologis, termasuk respons terhadap stresor[5], [7], [8]. Keunggulan teknologi EKG modern, seperti kemampuannya untuk melakukan perekaman data secara real-time, portabilitas melalui perangkat wearable, dan sifatnya yang non-invasif, menjadikannya sumber informasi yang kaya dan praktis[8], [9]. Potensi ini sangat relevan dalam pengembangan sistem klasifikasi tingkat stres otomatis yang dapat diintegrasikan dalam pemantauan kesehatan berkelanjutan, terutama dengan memanfaatkan kapabilitas machine learning.

Dalam domain kecerdasan buatan (Artificial Intelligence), khususnya machine learning, beragam algoritma klasifikasi telah dieksplorasi dan diterapkan untuk menganalisis data EKG yang kompleks guna mengidentifikasi pola-pola halus yang berkorelasi dengan

tingkat stres. Di antara spektrum algoritma yang tersedia, Decision Tree dan Naive Bayes sering menjadi pilihan awal karena kombinasi antara kesederhanaan konseptual, efisiensi komputasi, dan performa yang seringkali kompetitif[10]. Algoritma Decision Tree menawarkan keunggulan dalam hal interpretability, di mana model yang dihasilkan berupa struktur pohon keputusan yang intuitif dan mudah dipahami bahkan oleh non-ahli, sehingga memudahkan validasi klinis. Di sisi lain, algoritma Naive Bayes, yang berlandaskan pada teorema Bayes dengan asumsi fundamental mengenai independensi kondisional antar fitur, dikenal karena kecepatan pelatihannya yang superior dan kemampuannya memberikan hasil yang robust meskipun dengan dataset yang tidak terlalu besar[10], [11].

Meskipun demikian, perlu disadari bahwa tidak ada algoritma machine learning yang bersifat superior secara universal. Setiap algoritma memiliki karakteristik, prasyarat, kekuatan, dan keterbatasan inheren yang dapat mempengaruhi kinerjanya tergantung pada sifat data yang diolah. Data biologis seperti EKG inherently bersifat kompleks, seringkali mengandung noise akibat artefak gerakan atau interferensi elektromagnetik[7], serta menunjukkan variabilitas intra dan inter-individu yang tinggi. Faktor-faktor ini dapat secara signifikan mempengaruhi bagaimana sebuah algoritma memodelkan hubungan antara fitur EKG dan tingkat stres[8]. Oleh karena itu, sebuah analisis komparatif yang sistematis dan mendalam terhadap algoritma Decision Tree dan Naive Bayes menjadi krusial. Studi ini bertujuan untuk mengevaluasi secara cermat kinerja kedua algoritma tersebut dalam konteks spesifik klasifikasi tingkat stres menggunakan data EKG, dengan fokus pada metrik evaluasi kuantitatif seperti akurasi, presisi, recall, dan F1-score, serta mempertimbangkan aspek efisiensi komputasi (waktu pelatihan dan inferensi). Hasil dari analisis ini diharapkan dapat memberikan panduan berbasis bukti dalam memilih algoritma yang lebih optimal untuk aplikasi deteksi stres berbasis EKG.

1.2. Rumusan Masalah

Penelitian ini dilatarbelakangi oleh kebutuhan akan metode klasifikasi yang akurat dan efisien dalam mendeteksi tingkat stres berdasarkan data fisiologis dan kebiasaan individu. Dalam hal ini, algoritma pembelajaran mesin seperti Decision Tree dan Naive Bayes memiliki karakteristik masing-masing yang dapat dimanfaatkan untuk proses klasifikasi. Oleh karena itu, rumusan masalah yang diangkat dalam penelitian ini adalah bagaimana performa algoritma Decision Tree dan Naive Bayes dalam mengklasifikasikan tingkat stres berdasarkan data yang mencakup parameter seperti detak jantung, tekanan darah, kualitas tidur, dan aktivitas fisik, serta algoritma mana di antara keduanya yang menunjukkan hasil

klasifikasi paling optimal berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score.

1.3. Tujuan dan Manfaat Penelitian

1.3.1. Tujuan

Penelitian ini bertujuan untuk mengevaluasi dan membandingkan kinerja dua algoritma klasifikasi, yakni Decision Tree dan Naive Bayes, dalam mengidentifikasi tingkat stres berdasarkan data numerik dan kategorikal yang berhubungan dengan kondisi fisiologis individu. Melalui pengujian terhadap dataset yang memuat informasi seperti usia, aktivitas harian, kualitas tidur, serta parameter kesehatan seperti tekanan darah dan detak jantung, penelitian ini diharapkan dapat mengungkap seberapa efektif masing-masing algoritma dalam melakukan klasifikasi tingkat stres. Selain itu, penelitian ini juga bertujuan untuk menentukan algoritma yang memberikan hasil paling akurat dan stabil dalam konteks klasifikasi berbasis data kesehatan.

1.3.2. Manfaat

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Menambah pemahaman dalam bidang pembelajaran mesin, khususnya terkait efektivitas algoritma Decision Tree dan Naive Bayes dalam mendeteksi tingkat stres berdasarkan data fisiologis.
2. Menjadi referensi ilmiah bagi peneliti atau praktisi yang mengembangkan sistem pendeteksi stres berbasis data kesehatan dan kecerdasan buatan.
3. Memberikan dasar bagi pengembangan teknologi pemantauan kesehatan yang lebih presisi dan responsif terhadap kondisi pengguna melalui analisis data fisiologis.

1.4. Batasan Penelitian

Penelitian ini memiliki sejumlah batasan agar ruang lingkup kajian tetap fokus dan tidak melebar ke luar topik utama. Adapun batasan-batasan yang ditetapkan adalah sebagai berikut:

1. Penelitian ini hanya menggunakan data sekunder berupa hasil pengukuran dan informasi pribadi yang telah tersedia dalam bentuk dataset terstruktur. Data tersebut mencakup parameter seperti usia, jenis kelamin, profesi, durasi dan kualitas tidur, aktivitas fisik, indeks massa tubuh (BMI), tekanan darah, detak jantung, jumlah langkah harian, dan tingkat stres.

2. Data sinyal EKG mentah tidak digunakan dalam bentuk sinyal waktu atau grafik. Fokus penelitian ini adalah pada data numerik dan kategorikal yang dapat merepresentasikan kondisi fisiologis dan gaya hidup seseorang yang berhubungan dengan stres.
3. Algoritma klasifikasi yang diuji terbatas pada dua metode, yaitu Decision Tree dan Naive Bayes, tanpa membandingkannya dengan algoritma lain seperti SVM, KNN, atau Random Forest.
4. Variabel target (label) dalam penelitian ini adalah tingkat stres, yang awalnya berada dalam skala 1 hingga 10. Untuk keperluan klasifikasi, skala ini dapat dikategorikan ke dalam beberapa kelas seperti rendah, sedang, dan tinggi.
5. Penelitian ini tidak mencakup pengembangan sistem atau aplikasi berbasis real-time maupun integrasi dengan perangkat medis atau wearable. Fokus utama adalah pada analisis komparatif kinerja model klasifikasi terhadap data yang ada.

1.5. Keterbaruan

Terdapat beberapa penelitian sebelum nya yang memiliki keterkaitan dengan klasifikasi Kanker Paru-paru yang menunjukkan keterbaruan dari penelitian ini.

1. D. F. Astari, Y. H. Chrisnanto, Melina[12] mengungkapkan bahwa algoritma Random Forest mampu menghasilkan akurasi sebesar 93,65% dalam mengklasifikasikan tingkat stres, mengungguli algoritma C4.5 yang hanya mencapai 87,88%. Keunggulan Random Forest terletak pada mekanisme ensemble-nya, di mana setiap pohon keputusan dibangun dari kombinasi fitur yang berbeda-beda, sehingga meningkatkan ketahanan terhadap noise dalam data. Temuan ini menegaskan bahwa Random Forest merupakan pendekatan yang efektif dalam klasifikasi stres, terutama dalam konteks data tidur. Penelitian ini juga menyoroti pentingnya tahap praproses data dalam menentukan performa klasifikasi akhir.
2. S. S. A. Larasati, E. N. K. Dewi, B. H. Farhansyah, F. A. Bachtiar, F. Pradana[13] menyoroti bahwa pendekatan berbasis kondisi tidur memiliki potensi besar dalam mendeteksi tingkat stres, mengingat tidur merupakan aktivitas rutin yang menghasilkan data relatif stabil dan minim bias. Dalam studi ini, algoritma Decision Tree digunakan untuk membangun model klasifikasi dan berhasil mencapai akurasi sebesar 0,99, meskipun masih terdapat kasus salah klasifikasi. Model kemudian ditingkatkan dengan algoritma Random Forest, yang mampu memperbaiki kesalahan tersebut dan memberikan hasil sempurna dengan akurasi 1.0. Studi ini juga menunjukkan bahwa faktor-faktor seperti tingkat dengkur,

laju napas, gerakan anggota tubuh (termasuk gerakan mata saat tidur/REM), dan detak jantung cenderung berkorelasi positif dengan tingkat stres, sedangkan suhu tubuh, kadar oksigen darah, dan durasi tidur menunjukkan hubungan yang sebaliknya.

3. N. F. Mahing, A. L. Gunawan, A. F. Azhar Zen, F. A. Bachtiar, S. A. Wicaksono [14] mengevaluasi klasifikasi stres berbasis teks menggunakan algoritma Support Vector Machine (SVM) dan Random Forest yang dikombinasikan dengan berbagai metode transformasi data teks. Hasil penelitian menunjukkan bahwa model SVM dengan pendekatan TF-IDF memberikan kinerja terbaik dengan akurasi 84%, sementara model Random Forest menunjukkan performa tertinggi ketika menggunakan CountVectorizer, dengan akurasi 80%. Berdasarkan hasil tersebut, dapat disimpulkan bahwa SVM lebih unggul dalam klasifikasi teks stres, dengan keunggulan akurasi sekitar 3% dibandingkan Random Forest dalam konteks studi ini.
4. D. H. Depari, Y. Widiastiwi, M. M. Santoni[11] melakukan perbandingan performa antara algoritma Decision Tree, Naive Bayes, dan Random Forest untuk klasifikasi penyakit jantung. Hasil penelitian menunjukkan bahwa dari sisi akurasi, Random Forest unggul dengan akurasi 75%, diikuti oleh Decision Tree (72%) dan Naive Bayes (71%). Namun, aspek kecepatan eksekusi juga menjadi pertimbangan penting. Meskipun Random Forest menawarkan akurasi yang lebih tinggi, proses pelatihannya memakan waktu lebih lama karena membutuhkan pembangunan banyak pohon. Oleh karena itu, jika kecepatan dalam memperoleh hasil menjadi prioritas, maka Decision Tree dianggap lebih praktis karena waktu pemrosesan yang lebih cepat dengan akurasi yang masih kompetitif.